DOCUMENT RESUME

ED 423 258		TM 028 993

AUTHOR Roos, Linda L.; Wise, Steven L.; Finney, Sara J.

TITLE Comparing Restricted and Unrestricted Self-Adapted Testing as Alternatives to Computerized Adaptive Testing.

PUB DATE 1998-04-00

NOTE 22p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (San Diego, CA, April 13-17 1998).

PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.

DESCRIPTORS *Adaptive Testing; *College Students; Comparative Analysis; *Computer Assisted Testing; *Difficulty Level; Error of Measurement; Higher Education; Performance Factors; Selection; Test Anxiety; *Test Items

IDENTIFIERS *Restricted Self Adapted Testing

ABSTRACT

Previous studies have shown that, when administered a self-adapted test, a few examinees will choose item difficulty levels that are not well-matched to their proficiencies, resulting in high standard errors of proficiency estimation. This study investigated whether the previously observed effects of a self-adapted test--lower anxiety and higher test performance relative to a computerized adaptive test (CAT)--can be sustained while eliminating the high standard errors. A restricted self-adapted test (RS-AT) in which examinees were allowed to choose among a set of difficulty levels only in the region of their proficiency estimates was utilized in this study. Data were collected from 273 students in an introductory statistics class. The results show that while the RS-AT effectively controlled the standard errors of proficiency estimation, examinees receiving an RS-AT did not show higher mean proficiency or lower posttest state anxiety than examinees receiving a CAT. (Contains 3 tables and 15 references.) (SLD)

ED 423 258

Comparing Restricted and Unrestricted Self-Adapted Testing as Alternatives to Computerized Adaptive Testing

Linda L. Roos, Steven L. Wise, Sara J. Finney

University of Nebraska–Lincoln

TM028993

Paper presented at the 1998 annual meeting of the National Council on Measurement in Education, San Diego, CA

Abstract

Previous studies have shown that, when administered a self-adapted test, a few examinees will choose item difficulty levels that are not well matched to their proficiencies, resulting in high standard errors of proficiency estimation. This study investigated whether the previously observed effects of a self-adapted test—lower anxiety and higher test performance relative to a computerized adaptive test (CAT)—can be sustained while eliminating the high standard errors. A restricted self-adapted test (RS-AT) in which examinees were allowed to choose among a set of difficulty levels only in the region of their proficiency estimates was utilized in this study. The results showed that, while the RS-AT effectively controlled the standard errors of proficiency estimation, examinees receiving an RS-AT did not show higher mean proficiency or lower posttest state anxiety than examinees receiving a CAT.

Comparing Restricted and Unrestricted Self-Adapted Testing as Alternatives to Computerized Adaptive Testing

The development of Item Response Theory (IRT)—along with the proliferation of microcomputers—has led to the implementation of computerized adaptive testing in many settings. A computerized adaptive test (CAT) uses an algorithm to match item difficulty to examinee proficiency. Essentially, if an item is answered incorrectly then an easier item is administered; if an item is answered correctly then a more difficult item is administered. Recently, some variants of CATs have been developed, including the self-adapted test (S-AT[1]; Rocklin and O'Donnell, 1987). A S-AT allows an examinee to choose the difficulty level of each item from among a number of (typically six to eight) difficulty levels. After the desired number of items has been administered, an examinee is assigned a proficiency estimate that has been calculated using IRT-based scoring procedures.

There is evidence that a S-AT may be an attractive type of computer-based test. Research has shown that those examinees who were administered a self-adapted test (S-AT) obtained higher proficiency estimates than those administered a CAT (Wise et al., 1992; Vispoel & Coffman, 1994; Roos, Wise & Plake, 1997). Several studies have also shown that proficiency estimates obtained with self-adapted testing are less related to anxiety than those obtained with computerized adaptive testing (Roos et al., 1997; Vispoel & Coffman, 1994; Vispoel, Rocklin, & Wang, 1994; Vispoel, Wang, de la Torre, Bleiler, & Dings, 1992). Other studies have shown mean examinee state anxiety to be lower after completing a S-AT than a CAT (Wise et al., 1992; Roos et al., 1997). It appears that a S-AT has a positive influence on the anxiety and motivation levels of examinees that is likely attributable to examinees having increased perceived control (Wise, 1994). Many psychological studies have shown that, in a stressful situation, people who desire control and perceive that they have some control over the source of stress exhibit lower anxiety, increased motivation and improved performance on cognitive tasks (Perlmuter & Monty, 1977).

Although the CAT algorithm is designed to match item difficulty to examinee proficiency—and thereby minimize measurement error—examinees taking a S-AT are free to choose items from any available difficulty level. Although examinees have shown a tendency to choose difficulty levels that are reasonably well matched to their proficiency estimates (Wise, Plake, Johnson, Roos, 1992; Johnson, Roos, Wise & Plake, 1991), a few examinees choose items that are poorly matched. This results in the standard error associated with the S-AT proficiency estimate being higher than it would have been with a CAT. The possibility of proficiency estimates with large standard errors is a major liability of self-adapted testing.

In an effort to combine the benefits of self-adapted testing while preventing examinees from choosing items not well matched to their proficiency estimates, Wise, Kingsbury and Houser (1993) developed a restricted self-adapted test (RS-AT). Restricted self-adapted testing allows the examinee to choose from the subset of item difficulty levels that are most closely matched to his/her level of proficiency. For example, assume that the items have been divided into nine levels. Each time an examinee chooses an item, he/she is allowed to choose from among the five levels closest to the current proficiency estimate. Hence, an examinee with a very low estimate might be allowed to choose from levels 1-5, while an examinee with a moderate estimate might choose among levels 3-7, and a highly proficient examinee might choose among levels 5-9. This should provide examinees some control over item difficulty selection while preventing the choice of items that are poorly matched to their proficiency levels.

This study investigated the precision and effects of an RS-AT. There were three research questions: (a) Does an RS-AT effectively control the magnitude of error in proficiency estimates, relative to a S-AT? (b) How does the mean proficiency estimate from the RS-AT compare to that from a S-AT and a CAT? (c) How does the mean posttest anxiety from the RS-AT compare to that from a S-AT and a CAT? In essence, this is an investigation of whether a RS-AT can effectively control error like a CAT, while preserving the positive effects of a S-AT.

Method

Participants

The participants in this study were enrolled in several sections of an introductory statistics course at a large midwestern university. Data were collected from 273 examinees during the spring and summer academic sessions of 1997. The participants included approximately one-third graduate and about two-thirds undergraduate students; approximately one third were males and two-thirds were females.

Instruments

The primary instrument utilized in this study was a computerized algebra test designed to assess whether students possess the algebra skills necessary to be successful in an introductory statistics course. Each 25-item test was drawn from a pool of 144 four-option multiple choice items testing basic algebra skills. The pool was calibrated using a modified one-parameter IRT model that used a 0.20 common lower asymptote. Proficiency was estimated using maximum likelihood.

Three versions of the test were administered: CAT, S-AT and RS-AT. The CAT used a maximum information algorithm to determine which item should be administered to the examinee based on whether the examinee answered the previous items correctly or incorrectly. The instructions presented at the beginning of the test to those who were administered a CAT were:

> This 25-item test is intended to measure your level of proficiency in the types of mathematics skills that are needed for a course in introductory statistics. This test is different from most tests that you have taken. The items that you receive are chosen by the computer based on your performance. That is, every time you pass an item, you'll be given a more difficult item; every time you fail an item, you'll be given an easier item. Using this method, the computer will try to identify items that are reasonably matched to your algebra proficiency level. When calculating your score on this test, the computer will take into account the difficulty levels of the items you have received, and credit your answers accordingly.

The S-AT allowed examinees to choose the difficulty level of each item to be administered from among five levels of difficulty. The items within each difficulty level were randomly arranged and each examinee received the items from a difficulty level in the same order. The range of difficulty (b-parameters) for each of the difficulty levels were: level 1 (-5.359 to -1.390), level 2 (-1.389 to -0.666), level 3 (-0.649 to 0.0031), level 4 (0.0169 to 0.5343) and level 5 (0.5699 to 4.0077). The instructions presented to examinees who were administered the S-AT were:

> This 25-item test is intended to measure your level of proficiency in the types of mathematics skills that are needed for a course in introductory

> statistics. This test is different from most tests that you have taken. Before each test item is presented, you will choose how difficult you want the item to be. You will choose among five different levels of difficulty, ranging from level 1 (easier items) to level 5 (harder items).
>
> The higher the difficulty level of an item that you choose, the more credit you will receive if you pass the item. When calculating your score on this test, we will take into account the difficulty levels of the items you have chosen, and credit your answers accordingly.
>
> We recommend that you choose the hardest items that you think that you can answer correctly. You are, however, free to choose whatever item difficulty levels that you prefer. The items are weighted in such a way that it should not matter which items you have chosen—your final score should be about the same.

The RS-AT provided examinees with limited choice over the difficulty level of each item administered. The items were divided into nine difficulty levels and when making an item difficulty level selection, an examinee would have access to the five contiguous difficulty levels closest to his/her proficiency estimate. Because of the total number of items in the pool, each of the nine levels contained fewer than 25 items. The number of items contained in each level and the difficulty ranges were: level 1 (18; -5.359 to -1.726), level 2 (16; -1.6983 to -1.275), level 3 (15; -1.272 to -0.9022), level 4 (14; -0.831 to -0.536), level 5 (15; -0.5168 to -0.163), level 6 (14; -0.129 to 0.0732), level 7 (14; 0.0955 to 0.4572), level 8 (16; 0.472 to 0.8449) and level 9 (18; 1.0695 to 4.0077). If an examinee exhausted the items in a difficulty level, he or she was instructed to choose from another difficulty level. The instructions presented to RS-AT examinees differed from those presented to S-AT examinees only in the first paragraph:

> This 25-item test is intended to measure your level of proficiency in the types of mathematics skills that are needed for a course in introductory statistics. This test is different from most tests that you have taken. Before each test item is presented, you will have some control over its

> difficulty. Although the computer will try to identify items that are reasonably matched to your algebra proficiency level, you will be asked to choose the relative difficulty of each item. You will choose among five different levels of difficulty, ranging from level 1 (easier items) to level 5 (harder items).

In addition to the algebra test, three other instruments were administered to examinees, each using a paper and pencil format. The Test Anxiety Inventory (TAI; Spielberger, 1980) measured examinee test anxiety. The Desire for Control on Examinations scale (DCE; Wise, Roos, Leland, Oats, & McCrann, 1996) measured the desire for control expressed by examinees in a testing context. The State Anxiety Scale (Spielberger, Gorsuch, & Lushene, 1970) was administered immediately before and after the algebra test to measure situation-specific anxiety of the examinees.

Procedure

During the first class session, participants supplied demographic information, completed the TAI and the DCE, and signed up for an algebra test administration time. The participants were informed that those who did not score above a pre-determined cutoff on the algebra test would be required to attend an one hour algebra review session held early in the term.

The algebra test was administered in a room containing 12 Dell Pentium microcomputers running MicroCAT™ (Assessment Systems, 19) software[2]. Examinees were randomly assigned to one of the three test conditions (CAT, S-AT or RS-AT), asked to read and sign a consent form, and complete the State Anxiety Scale. Next, the testing software presented the appropriate instructions describing the assigned testing procedure, and then administered the algebra test. Scratch paper and pencils were provided and calculators were not allowed. No time limit was imposed during testing. Upon completion of the algebra test, the examinees were again asked to complete the State Anxiety Scale. Then, the examinees were asked to respond to several questions that were presented electronically. For the first question, which asked, "How clear were the instructions given at the beginning of the test?", examinees responded using a five-point scale ranging from not at all clear to very clear. The second question asked "How much control did you feel you had over your test performance?", using a five-point scale ranging from no control to a great deal of control. Examinees in the S-AT and RS-AT conditions responded to a third question which asked, "To what degree do you feel that you were able to control the difficulty of your test?", using a 5-point scale of responses ranging from no control to a great deal of control. Finally, the examinees were informed whether they were required to attend a review session.

Data Analysis

The first research question concerning relative measurement error among the test types was evaluated by inspection of the minimum, median and maximum standard errors of proficiency estimate for each condition. Because the standard error distributions were likely to be skewed, Mann-Whitney *U* tests were used to evaluate the significance of the differences in the standard errors between each pair of test types. The second research question concerning relative mean proficiency among the test types was evaluated using an analysis of covariance (ANCOVA) with test type as the independent variable, estimated proficiency as the dependent variable and number of years since last algebra course as the covariate. The third research question concerned differences among the test types in mean posttest anxiety. An ANCOVA was performed using test type as the independent variable, posttest state anxiety as the dependent variable and pretest state anxiety as the covariate.

Results

Table 1 presents the minimum, median and maximum standard error of proficiency estimate for each of the testing conditions. As expected, the minimum and median standard errors for the RS-AT were very similar to that observed for the CAT. The maximum standard error, however, was much higher for the S-AT than for the other two test types. Mann-Whitney *U* tests showed that the S-AT differed significantly from both the CAT ($z = -4.76$, $p < .001$) and the RS-AT ($z = -3.12$, $p = .002$) but the CAT and the RS-AT did not differ significantly from each other ($z = -1.60$, $p = .110$). Large standard errors occurred with the S-AT because several examinees chose items poorly matched to their proficiency levels.

Table 1

Descriptive Statistics for Standard Error of Proficiency Estimation, By Experimental Condition

	Experimental Condition		
Standard Error	S-AT	RS-AT	CAT
Minimum	0.08	0.08	0.09
Median	0.12	0.11	0.10
Maximum	24.83	0.32	0.65

Table 2 shows the means and standard deviations, by test type, for a number of outcome variables including estimated proficiency and posttest state anxiety. The adjusted means from the ANCOVAs for both estimated proficiency and posttest state anxiety are shown in Table 3. Regarding estimated proficiency, no significant differences were found among the test types. The analysis of posttest state anxiety revealed significant differences among the test types. Tukey follow-ups (using a 0.05 familywise significance level) showed that the S-AT yielded posttest anxiety levels

Table 2

Means and Standard Deviations of Study Outcome Variables, by Experimental Condition

	Experimental Condition					
	S-AT (n = 93)		RS-AT (n = 86)		CAT (n = 94)	
Variable	Mean	SD	Mean	SD	Mean	SD
Estimated Proficiency	0.04	1.34	-0.14	1.15	-0.06	1.31
Posttest State Anxiety	38.76	12.50	40.80	11.39	41.12	11.60
Number of Items Passed	17.45	3.99	16.33	2.78	16.74	3.00
Average Item Difficulty	-0.36	0.89	-0.35	1.05	-0.35	1.05
Average Item Targeting	-0.41	0.95	-0.21	0.46	-0.29	0.56
Clarity of Instructions	4.42	0.83	4.66	0.79	4.59	0.74
Control Over Performance	3.74	1.11	3.59	1.09	3.76	1.08
Control Over Difficulty	4.20	0.97	3.98	1.20		

Table 3

Adjusted Means for Estimated Proficiency and Posttest State Anxiety

	Experimental Condition		
Variable	S-AT	RS-AT	CAT
Estimated Proficiency	0.08	-0.18	-0.05
Posttest State Anxiety	37.97	41.64	40.95

that were significantly lower than either the CAT or the RS-AT, which did not significantly differ from each other.

For both estimated proficiency and posttest state anxiety, the magnitude of the effects found in the differences between CAT and S-AT are similar to that observed in previous studies using a similar item pool and examinees possessing similar demographics. The RS-AT and CAT are similar not only in observed standard error of proficiency but also in proficiency estimates.

To gain insight regarding why poorly matched item difficulty levels were chosen, the characteristics of examinees who exhibited high standard errors were studied. There were three examinees whose standard errors exceeded 0.7; each of these examinees was (a) administered a S-AT, (b) completed the test in slightly less than the average time for those administered a S-AT, and (c) reported recently completing an algebra course. Beyond these variables, however, the cases were markedly different.

The first examinee (Examinee A), was a male who consistently chose the third difficulty level and answered all of the items correctly. He exhibited low pretest anxiety and moderate desire for control, and indicated that he felt that he was able to control both his test performance and the difficulty of his test. This information suggests that Examinee A was never engaged in the process of taking the S-AT. He may not have been motivated to excel on the test, possibly because he was confident of exceeding the standard for acceptable performance.

The second examinee (Examinee B), was a female who began her test by choosing and passing four items from the third difficulty level. She then attempted and failed an item from the fourth difficulty level. At this point, her selection behavior changed dramatically. For 18 of the remaining 20 items, Examinee B chose the first difficulty level, answering only nine of them correctly. She exhibited high pretest anxiety, moderate desire for control and indicated that she felt that she was able to control both her test performance and the difficulty of her test. It appears that, after some early success on the test, Examinee B disengaged from the task when she encountered failure. That is, her performance on the moderately difficult items from the early part of her test suggests that she was fairly proficient, whereas her poor performance on the remainder of the test was consistent with an examinee of low proficiency. It was this inconsistency in her testing session that produced her high standard error.

The third examinee (Examinee C), was a female who reported high pretest anxiety and high desire for control. On the first three items of her test, Examinee C failed items from the third, second, and first difficulty levels, respectively. The last

22 items of her test were primarily chosen from the fourth and fifth difficulty levels, and she answered nine correctly. Examinee C indicated that she felt that she was able to control neither her test performance nor the difficulty of her test. It appears that she attempted to escape the stress of the test through selection of inordinately difficult items. That is, after her early incorrect answers, she ensured failure on the test by subsequently choosing items that she was sure to fail thus rendering inevitable the outcome of the test.

Although the characterizations of these examinees' reactions to the testing experience are admittedly speculative, it appears clear that they behaved in distinctly different ways. It is therefore likely that a variety of examinees could potentially choose poorly-matched difficulty levels during a S-AT.

Discussion

Since its introduction a decade ago, self-adapted testing has represented an intriguing alternative to computerized adaptive testing. It has shown promise as a testing procedure that can decrease the impact of test anxiety on examinee test performance. One of its key limitations, however, is that examinees can attain proficiency estimates with unacceptably high standard errors through selection of difficulty levels that are poorly matched to proficiency. Until this limitation is overcome, it is unlikely that self-adapted testing will be adopted by an operational testing program.

Our results indicate that, although the RS-AT was effective in controlling the large standard errors that had previously been observed with S-ATs, its effect on examinees more closely resembled a CAT than a S-AT. Positive effects that have been observed with the S-AT—higher mean estimated proficiency and lower posttest state anxiety—were not realized with the RS-AT. Rather, mean proficiency and anxiety for examinees receiving the RS-AT were more similar to that observed with those receiving the CAT. Although the problem of large standard errors was successfully addressed by the R-SAT used in this study, the positive effects of self-adapted testing were absent. We were, therefore, ultimately unsuccessful in achieving our general goal of developing a self-adapted test that alleviated the effects of test anxiety while controlling standard errors.

Further study of the basic RS-AT procedure is warranted. Although the results of this study are not encouraging, modifications to the procedures used in this study could be explored. Some issues to consider for future studies involving RS-AT include (a) the clarity of instructions presented at the beginning of the test, (b) the number of difficulty levels presented to examinees as well as the total number of difficulty levels, (c) the labeling of strata, and (d) training regarding the RS-AT procedure. It is possible that, in the current study, examinees may not have understood, for example, that the third difficulty level choice that appeared on their computer screen could correspond to different levels of absolute difficulty, depending upon their current proficiency estimate. If that were the case, and a given item difficulty level choice did not always correspond to the same absolute level of difficulty, confused examinees may have doubted the degree to which they were actually being permitted to control item difficulty. Thus, if examinees were confused regarding the instructions, then the credibility of the R-SAT procedure would be undermined, and it would not be surprising that their mean proficiency and anxiety resembled those obtained with the CAT—in which examinee control over difficulty was also not provided.

This study investigated only one configuration of the RS-AT procedure, in which examinees were presented item difficulty choices from among five out of nine difficulty levels. It is not clear if the R-SAT would have yielded different results if a different number of choices and/or total difficulty levels had been used.

In the present study, the RS-AT difficulty level choice screens always presented choices among difficulty levels one through five, regardless of the absolute difficulty levels. To alleviate any confusion between the difficulty levels presented on the computer screen and the absolute levels of difficulty available, strata could be labeled to indicate to which strata the examinees have access. That is, if an examinee has access to the third through seventh strata, the difficulty level choice screen would indicate that. This strategy would not be possible if review were allowed because an examinee with knowledge of the CAT algorithm may be able to tell which items had been answered correctly based on the strata presented.

It is also possible that the results for the RS-AT could change if more extensive training regarding the testing procedure was provided. Both the S-AT and RS-AT represent novel testing situations for nearly all examinees and it is not clear how examinees would perform if they were more accustomed to these testing formats. Also, it is not clear how examinees would perform on either S-AT or RS-AT if the tests were administered in a higher-stakes testing environment. In a high stakes testing situation, presumably there would be additional training regarding the testing procedure; if examinees better understood the amount of control possible with the RS-AT, the results may differ.

It remains unclear whether a self-adapted testing procedure can provide examinees credible choice over item difficulty, while preventing them from making poorly matched choices. To the extent that the perceived control hypothesis is correct (Wise, 1994), then the effects of a self-adapted test on test performance are dependent on an examinee's perception of control. Hence, the ideal self-adapted test should provide control, but not too much control. Exploration of testing procedures that attempt to balance these psychological and psychometric demands should continue.

Footnotes

[1] Historically, self-adapted testing has been referred to as SAT. Henceforth, it will be referred to as S-AT to alleviate confusion with the Scholastic Achievement Test.

[2] For information regarding MicroCAT code for both self-adapted and restricted self-adapted tests, consult Roos, Wise, Yoes, & Rocklin, (1996).

Author Note

The authors wish to express thanks to Craig Enders for assistance in the data collection portion of this study.

Correspondence regarding this manuscript should be addressed Linda Roos; 501 Building, Room 118; University of NE-Lincoln; Lincoln, NE 68588-0203. Electronic mail may be sent to lroos@unl.edu.

References

Assessment Systems (1994). *User's Manual for the MicroCAT™ Testing System*. (3rd Edition, Revised). St. Paul, MN: Assessment Systems Corporation.

Johnson, P. L., Roos, L. L., Wise, S. L., & Plake, B. S. (1991). Correlates of examinee item choice behavior in self-adapted testing. *Mid-Western Educational Researcher, 4*, 25-29.

Perlmuter, L. C., & Monty, R. A. (1977) The importance of perceived control: Fact or fantasy? *American Scientist*, 65, 759-765.

Rocklin, T., & O'Donnell, A. M. (1987). Self-adapted testing: A performance-improving variant of computerized adaptive testing. *Journal of Educational Psychology, 79*, 315-319.

Roos, L. L., Wise S. L., & Plake, B. S. (1997). The role of item feedback in self-adapted testing. *Educational and Psychological Measuremen*t, 57, 85-98.

Roos, L. L., Wise, S. L., Yoes, M. E., & Rocklin, T. R. (1996). Conducting self-adapted testing using MicroCAT. *Educational and Psychological Measurement, 56*, 821-827.

Spielberger, C. D. (1980). *Preliminary professional manual for the Test Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Vispoel, W. P., Rocklin, T. R., & Wang, T. (1994). Individual differences and test administration procedures: A comparison of fixed-item, computerized-adaptive, and self-adapted testing. *Applied Measurement in Education, 7*, 53-79.

Vispoel, W. P., & Coffman, D. D. (1994). Computerized-adaptive and self-adapted music-listening tests: Psychometric features and motivational benefits. *Applied Measurement in Education, 7,* 25-51.

Vispoel, W. P., Wang, T., de la Torre, R., Bleiler, T., & Dings, J. (1992, April). *How review options and administration modes influence scores on computerized vocabulary tests.* Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.

Wise, S. L., (1994). Understanding self-adapted testing: The perceived control hypothesis. *Applied Measurement in Education, 7,* 15-24.

Wise, S. L., Kingsbury, G. G., & Houser, R. L. (1993, April). *An investigation of restricted self-adapted testing.* Paper presented at the Annual Meeting of the National Council on Measurement in Education, Atlanta, GA.

Wise, S. L., Plake, B. S., Johnson, P. L., & Roos, L. L. (1992). A comparison of self-adapted and computerized adaptive tests. *Journal of Educational Measurement, 29,* 329-339.

Wise, S. L., Roos, L. L., Leland, V. L., Oats, R. G., & McCrann, T. O. (1996). The development and validation of a scale measuring desire for control on examinations. *Educational and Psychological Measurement, 56,* 710-718.

ERIC®

TM028993

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Comparing Restricted and Unrestricted Self-Adapted Testing as Alternatives to Computerized Adaptive Testing	
Author(s): Linda L Roos, Steven L. Wise, Sara J. Finney	
Corporate Source:	Publication Date: 4/1998 (NCME)

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 1	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2A	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2B
Level 1 ☑	Level 2A ☐	Level 2B ☐
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, → please

Signature: Linda L Roos	Printed Name/Position/Title: Linda Roos Mgr, User Ed, Information Services	
Organization/Address: U of NE-Lincoln Lincoln, NE 68588-0203	Telephone: (402) 472-5227	FAX: (402) 472-8486
	E-Mail Address: lroos@unl.edu	Date: 5/21/98

(over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

THE UNIVERSITY OF MARYLAND
ERIC CLEARINGHOUSE ON ASSESSMENT AND EVALUATION
1129 SHRIVER LAB, CAMPUS DRIVE
COLLEGE PARK, MD 20742-5701
Attn: Acquisitions

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
1100 West Street, 2nd Floor
Laurel, Maryland 20707-3598

Telephone: 301-497-4080
Toll Free: 800-799-3742
FAX: 301-953-0263
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com